Amendment to the Claims:

This listing of claims will replace all prior versions, and listing of claims in the application.

1. (Original) A speech recognition system comprising an acoustic detector for detecting speech utterances of a speaker; a visual detector for detecting at least one facial characteristic associated with speech utterances of the speaker; a processing arrangement connected to be responsive to the acoustic and visual detectors for deriving a signal having first and second values respectively indicative of the speaker making and not making speech utterances such that the first value is derived only in response to the acoustic detector detecting a finite, nonzero acoustic response while the visual detector detects at least one facial characteristic associated with speech utterances of the speaker; and a speech recognizer for deriving an output indicative of the speech utterances as detected only by the acoustic detector, the speech recognizer being connected to be responsive to the acoustic detector only while the signal has the first value.

2. (Original) The speech recognition system of claim 1 wherein the processing arrangement causes the signal to have the second value in response to any of: (a) the acoustic detector not detecting a finite, nonzero acoustic response while the visual detector does not detect speech utterances of the speaker, (b) the acoustic detector detecting a finite, nonzero acoustic response while the visual detector does not detect speech utterances of the speaker, and (c) the acoustic detector not

detecting a finite, nonzero acoustic response while the visual detector detects speech utterances of the speaker.

3. (Original) The speech recognition system of claim 2 wherein the processing arrangement includes a delay arrangement for assuring that the beginning of each speech utterance is coupled to the speech recognizer.

4. (Original) The speech recognition system of claim 3 wherein the delay arrangement includes a memory element connected to be responsive to the acoustic detector, the memory element including a plurality of stages for storing sequential segments of the output of the acoustic detector, the delay arrangement being such that the contents of the memory element stage storing the beginning of a speech utterance are initially coupled to the speech recognizer.

5. (Original) The speech recognition system of claim 4 wherein the memory element includes a ring buffer.

6. (Original) The speech recognition system of claim 1 wherein the processing arrangement includes a delay arrangement for assuring that the beginning of each speech utterance is coupled to the speech recognizer.

7. (Original) The speech recognition system of claim 6 wherein the delay

arrangement includes a memory element connected to be responsive to the acoustic detector, the memory element including a plurality of stages for storing sequential segments of the output of the acoustic detector, the delay arrangement being such that the contents of the memory element stage storing the beginning of a speech utterance are initially coupled to the speech recognizer.

8. (Original) The speech recognition system of claim 7 wherein the delay arrangement is arranged for assuring that upon the completion of each speech utterance the acoustic detector is decoupled from the speech recognizer.

9. (Original) The speech recognition system of claim 1 wherein the processing arrangement includes a delay arrangement for assuring that upon the completion of each speech utterance the acoustic detector is decoupled from the speech recognizer.

10. (Original) The speech recognition system of claim 8 wherein the delay arrangement includes a memory element connected to be responsive to the acoustic detector, the memory element including a plurality of stages for storing sequential segments of the output of the acoustic detector, the delay arrangement being such that the contents of the memory element stage storing acoustic energy associated with the acoustic detector and which occurs upon completion of each speech utterance is prevented from being coupled to the speech recognizer.

11. (Original) The speech recognition system of claim 9 wherein the memory element includes a ring buffer.

12. (Original) The speech recognition system of claim 1 wherein the processing arrangement includes a face recognizer connected to be responsive to the visual detector.

13. (Original) The speech recognition system of claim 12 wherein the face recognizer is arranged for enabling the signal to have the first value only in response to the face of the speaker being at a predetermined orientation relative to the visual detector.

14. (Original) The speech recognition system of claim 13 wherein the face recognizer is arranged for: (1) detecting and distinguishing the faces of a plurality of speakers, and (2) enabling the signal to have the first value only in response to the speaker having a recognized face.

15. (Original) The speech recognition system of claim 14 wherein the processing arrangement includes a speaker identity recognizer connected to be responsive to the acoustic detector, the speaker identity recognizer being arranged for: (1) detecting and distinguishing speech patterns of a plurality of speakers, and (2)

enabling the signal to have the first value only in response to the speaker having a recognized speech pattern.

16. (Original) The speech recognition system of claim 15 wherein the processing arrangement is arranged for causing the signal to have the first value only in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

17. (Original) The speech recognition system of claim 1 wherein the processing arrangement includes a face recognizer connected to be responsive to the visual detector and a speaker identity recognizer connected to be responsive to the acoustic detector, the face recognizer being arranged for detecting and distinguishing the faces of a plurality of speakers, the speaker identity recognizer being arranged for detecting and distinguishing speech patterns of a plurality of speakers, the processing arrangement being arranged for enabling the signal to have the first value only in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

18. (Original) A method of recognizing speech utterances of a speaker with an automatic speech recognizer only responsive to acoustic speech utterances of the speaker comprising detecting acoustic energy having a spectrum associated with speech utterances, detecting at least one facial characteristic associated with

speech utterances of the speaker, and activating the automatic speech recognizer only in response to the detected acoustic energy having a spectrum associated with speech utterances while the at least one facial characteristic associated with speech utterances of the speaker is occurring.

19. (Original) The method of claim 18 further comprising preventing activation of the automatic speech recognizer in response to any of: (a) no acoustic energy having a spectrum associated with speech utterances being detected while no facial characteristic associated with speech utterances of the speaker is detected, (b) acoustic energy having a spectrum associated with speech utterances being detected while no facial characteristic associated with speech utterances of the speaker is detected, and (c) no acoustic energy having a spectrum associated with speech utterances being detected while at least one facial characteristic associated with speech utterances of the speaker is detected.

20. (Original) The method of claim 18 further comprising assuring that the beginning of each speech utterance is coupled to the speech recognizer.

21. (Original) The method of claim 20 wherein the beginning of each speech utterance is assuredly coupled to the speech recognizer by: (a) delaying the speech utterance, (b) recognizing the beginning of each speech utterance, and (c) responding to the recognized beginning of each speech utterance to couple the

delayed speech utterance associated with the beginning of each speech utterance to the speech recognizer and thereafter sequentially coupling the remaining delayed speech utterances to the speech recognizer.

22. (Original) The method of claim 18 further comprising assuring that no detected acoustic energy is coupled to the speech recognizer upon the completion of a speech utterance.

23. (Original) The method of claim 22 wherein assurance that no detected acoustic energy is coupled to the speech recognizer upon the completion of a speech utterance is provided by: (a) delaying the acoustic energy associated with the speech utterance, (b) recognizing the completion of each speech utterance, and (c) responding to the recognized completion of each speech utterance to decouple delayed acoustic energy occurring after the completion of each speech utterance from the speech recognizer.

24. (Original) The method of claim 18 wherein the at least one facial characteristic indicates the face of the speaker has a predetermined orientation relative to a detector involved in the step of detecting the at least one facial characteristic.

25. (Original) The method of claim 24 further comprising distinguishing the face

of the speaker from a plurality of speakers, distinguishing the speech pattern of the speaker from a plurality of speakers, and activating the automatic speech recognizer only in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

26. (Original) The method of claim 18 further comprising distinguishing the face of the speaker from a plurality of speakers, distinguishing the speech pattern of the speaker from a plurality of speakers, and activating the automatic speech recognizer only in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

27. (Original) The method of claim 26 further including storing: (1) images of the faces of a plurality of speakers, and (2) the speech patterns of the same plurality of speakers during at least one training period; and performing the distinguishing steps by comparing the stored images and speech patterns with images of the face of the speaker and the speech pattern of the speaker.